Ross River virus and the necessity of multi-scale, eco-epidemiological analyses

Keywords: multi-scale, epidemiology, spatial, arbovirus, zoonotic, ecology, ecoepidemiology

Authors and affiliations (in order of acknowledgement):

1) Emily J. Flies
School of Pharmacy and Medical Sciences, University of South Australia, GPO Box 2471 Adelaide SA 5001, Australia

School of Biological Sciences, University of Tasmania, Private Bag 55, Hobart Tas 7001, Australia
Fax: +61 3 6226 7809 Telephone: +61 3 6226 2125 Email: Emily.Flies@utas.edu.au

2) Philip Weinstein
School of Biological Sciences, Adelaide University, Molecular Life Sciences Ground Level, North Terrace, Adelaide SA 5005, Australia

3) Sharolyn J. Anderson
School of Natural and Built Environments, University of South Australia, Mawson Lakes SA 5095, Australia

4) Iain Koolhof
Menzies Institute for Medical Research, University of Tasmania, 17 Liverpool St, Hobart TAS 7000, Australia

5) Johannes Foufopoulos
School of Sustainability and the Environment, University of Michigan, 440 Church St, Ann Arbor, MI 48109, USA

6) Craig R. Williams
University of South Australia, School of Pharmacy and Medical Sciences, GPO Box 2471, Adelaide SA 5001, Australia

summary: We find that the spatial scale/aggregation of an analysis influences the apparent importance of ecological drivers of arboviral (Ross River virus) disease; we urge future epidemiological studies to include multiple spatial scales for a more complete picture of disease drivers.

Background: Zoonotic vector-borne disease prevalence is affected by vector, human and reservoir host factors, which are influenced by habitat and climate; these five components interact on microhabitat to landscape scales but are often analyzed at a single spatial scale.

Methods: We present an information theoretic, multi-scale, multiple regression analysis of the ecological drivers of Ross River virus. We analyze the spatial pattern of 20 years of Ross River virus infections from South Australia (1992-2012; n = 5,261) using variables across these five components of disease ecology at three spatial scales.

Results: We found that covariate importance depended on the spatial scale of the analysis; some biotic variables were more important at fine scales and some abiotic variables were more important at coarser spatial scales. The urban score of an area was most predictive of infections and mosquito variables did not improve the explanatory power of these models.

Conclusions: Through this multi-scale analysis, we identified novel drivers of the spatial distribution of disease and recommend public health interventions. Our results underline that single-scale analyses may paint an incomplete picture of disease drivers, potentially creating a major flaw in epidemiological analyses. Multi-scale, ecological analyses are needed to better understand infectious disease transmission.

Introduction

Each year more than a billion people are infected and over a million people die from vector-borne pathogens [1]. Understanding the complex, ecologically-driven transmission patterns that ultimately result in human infections can shed light on previously unrecognized risk factors and

provide potential targets for breaking the transmission cycle. Thus, better knowledge of transmission ecology can lead to improved control, mitigation and education strategies.

Zoonotic arboviruses can only persist within a vector, reservoir, or human host and human infections and outbreaks occur at the ecological nexus of these three organisms. The survival, abundance and behavior of the hosts, vectors and arboviruses, are shaped by locally available habitat and climate; we conceptualize the components of zoonotic arbovirus transmission ecology in Figure 1. Mathematical models are often used to explain mosquito-borne disease patterns using data about human hosts (e.g. population density, infectious status) and vectors (e.g. biting rate, longevity). But these models rarely consider potentially important climate, habitat, animal host and spatial data [2]. Eco-epidemiological analyses, in contrast, can examine the degree to which these five components (vectors, animal hosts, humans, habitat, and climate) relate to the distribution of arboviral infections [3]. Analyses conducted at coarser scales (e.g. global, national) can reveal ecological requirements for a pathogen. Finer scale analyses (e.g. regional, local) identify the disease dynamics of a specific area and can indicate where interventions will be most effective.

The spatial scale of an eco-epidemiological analysis is often determined ad-hoc, based on data availability, political boundaries and study objectives. The spatial (or temporal) scales at which disease data are analyzed can impact the appearance of disease patterns and therefore the prevention and control decision-making [4,5]. Scale can also influence the correlates of disease; biotic factors (e.g. species richness) are typically more important at fine spatial scales while climatic variables dictate disease dynamics at coarser scales [6].

Because spatial scale can determine the outcome of the analysis, single-scale analyses may reveal an incomplete or incorrect picture of the ecological drivers of transmission [6] and lead to ineffective disease mitigation efforts [4].

One zoonotic infection that, due to its complex ecology, particularly lends itself to eco-epidemiological analysis is the Australian mosquito-borne virus, Ross River virus (RRV; *Togaviridae, Alphavirus*). Ross River virus is the most common and wide-spread vector-borne disease in Australia (1451-9554 cases nationally; [7]). The disease caused by RRV is characterized by polyarthralgia, rash and fatigue which can last for weeks or months [8]. Available treatment for the disease is symptomatic only, so preventing infection through mosquito control, habitat modification and education remain the most effective means of disease reduction.

Ross River virus has a complex ecology [9]; it has been isolated from over 40 species of mosquitoes and can amplify in at least 18 animal host species [10]. Analyses to predict and explain RRV disease outbreaks have been conducted throughout Australia; weather patterns (particularly temperature and rainfall) consistently explain temporal patterns of disease [11].

However, much less is known about the factors driving the spatial patterns of RRV disease. The few spatial analyses that have been conducted have found disease distribution to be correlated with socioeconomic, animal host, habitat, climatic and/or mosquito variables [11–17]. One particularly thorough study incorporated at least one variable from each of the five categories in Fig. 1 [18] but no study has explored whether or how the importance of these components is affected by the spatial scale of the analysis.

This paper aims to a) explore the effect of spatial scale on covariate importance, b) understand the competing importance of transmission ecology components on RRV disease prevalence (Fig. 1), and c) determine whether mosquito surveillance data improve explanatory models that include covariates from the other four arboviral ecology components. By achieving the goals set out here, our study improves understanding of RRV ecology, thereby potentially leading to more effective disease control and prevention. Furthermore, because the five transmission components and the question of the appropriate spatial scale of analyses are relevant for all zoonotic arboviruses, our findings can provide insight into the spatial ecology of other mosquito-borne diseases.

Methods

Study location and state geographical divisions

This study was conducted in South Australia, the fourth largest and fifth most populous state in Australia (Figure 2). Three geographical divisions were used for this analysis (Table 1), based on government shapefiles acquired from the Australian Bureau of Statistics [19]. More information on geographical divisions and other methods are provided in the Supplementary Materials (SM).

Human infection data

Arbovirus infection data were acquired from the South Australian Department of Health's Notifiable Infectious Disease Surveillance System database [20]. Data from cases of RRV infection notified between 1 January 1992 and 31 December 2012 (n = 5,261) were provided to the authors de-identified (Human ethics approval from UniSA: 0000030917 and SA Health: HREC/13/SAH/05). Patient-reported location of infection was used as the geolocating field for

this analysis since prior analysis and additional analyses here (Table S1) found this to be a more accurate geolocator than the typically used patient place of residence [20].

Covariates

Variables that were considered a priori to influence RRV distribution (Table 2) were collated for this analysis from various datasets and sources (see Table S2 for data sources and citations for a priori hypotheses). Figure 1 identifies the 16 selected covariates and how their classification in the transmission ecology was conceptualized. Critical details are given below and additional information (including covariate abbreviation, description, data source and citations supporting the a priori hypotheses) is provided in Table S2.

Previous studies have identified outdoor recreation as a potential risk factor for RRV infection [20,21]; however the relationship has not been quantitatively tested. As a novel proxy for the emphasis on outdoor recreation for a given area, the number of caravan parks per person was calculated and used in this study.

The urban score covariate is based on the Global Human Settlement Layer (GHSL), which essentially places each area on an urban-rural gradient where higher scores mean a more urban environment [22].

Mosquito variables were calculated from a 17-year dataset of mosquito surveillance covering 8031 trap nights from 164 trap locations [23–25]. For each trap location (Fig. 2 insert), the maximum number of individuals from each species ever caught in a single night at the given location was determined. The maximum rather than the average number of mosquitoes at a location was used to account for the epidemic pattern of RRV: the ability of a location to support

outbreak mosquito populations was deemed more important than average mosquito abundance. There were not mosquito data available for each geographic division so the sample size of the "mosquito data subset" for each geographic division is listed in Table 1. Further details of covariate calculations are available in SM.

The Murray River and the coast represent the main water bodies in the state and major breeding grounds for the three principal vectors (*Culex annulirostris* inland and *Aedes camptorhynchus* and *Aedes vigilax* coastally).

Species richness was included because of its potential to modify disease dynamics by diluting or amplifying transmission [26].

Statistical models

The statistical computing platform R [27] was used for all data processing and analysis. Due to previously observed problems with interpreting standardized morbidity ratios [20], count data were used as the dependent variable for these analyses and the expected number of cases per geographic division was included as a covariate to account for different expected infection rates (based on age-standardized population in each geographic division; further explanation in SM). All covariates were standardized to z-scores (subtracting the mean and dividing by one standard deviation) to reduce the effects of collinearity and to allow for simple comparisons of effect sizes across predictor variables.

Generalized additive models with a negative binomial distribution were used to model relationships at all scales (model output in Table S3). Individual variables were compared for their importance using the change in AIC or percent deviance explained when a variable was

singly removed from the saturated model (Δ AIC or Δ % deviance explained, respectively). See SM for additional details of model fit and output and tests for collinearity and residual autocorrelation (Tables S3-S6, Figure S1).

There was substantial variation in the size of the polygons used for each spatial scale (Table 1). To test whether this variation impacted our results, we attempted the analysis with the data converted to a consistently-sized grid (raster). This approach necessitated disaggregating the data to a false resolution and was deemed inappropriate; details and results are in Table S7. Instead, the area of the polygons was added as a covariate and tested for significance (details in SM).

Results

Spatial distribution of human infections

Ross River virus disease incidence was largely concentrated along the Murray River, as were the highest densities of vector mosquitoes (Fig. 2 map insert). The data also suggest relatively high disease incidence in the very large areas to the north-east of the state.

Importance of scales

Table 3 shows the explanatory power of the individual variables at each of the three scales, excluding variables removed due to multicollinearity (Tables S4 & S5). The overall percent deviance explained was highest for the post codes spatial scale (53.0%) compared to the finer and coarser scales of suburb and region (40.2% and 45.3% percent deviance explained, respectively).

For the two finer scales, the expected number of cases was the most important variable (calculated using the average number of RRV infections per person per age group for the state and the age-stratified population of each suburb; see SM and [20]), but at the coarsest scale of aggregation (region), the expected number of cases became negatively associated with disease. The other main driver for the observed number of cases was the urban score of the area. Urban score was negatively associated with number of cases and was the second most important covariate at the two finer scales and the most important variable at the coarsest (regional) scale (Fig. 3).

Other variables that influenced the models (Δ AIC reduced by more than 2 when variable was removed from saturated model) were socioeconomic status (negative relationship only at finest scale), distance to coast (positive relationship only at coarsest scale), bird biodiversity (negatively related at two finest scales), caravan parks per person (positively related at the two finer scales), elevation (negatively related at all scales), and mammal biodiversity (positively related only at finest scale).

The distance to the Murray River (DistToMurray) was an important component of the saturated model: on every spatial scale it had a strong (greater than 3) influence on model AIC and, when removed, reduced the percent deviance explained by 1-2.6%. However, the DistToMurray single model coefficient was zero or almost zero for at all scales, indicating that while this covariate does not explain much on its own, it modifies the relationship among other covariates, making it an important saturated model component. Most notably, maximum and minimum temperature are unimportant predictors at all scales when DistToMurray is included in the model but appear

to be important at all scales when DistToMurray is not included in the model because these variables are strongly modified by their proximity to the Murray River.

The value of mosquito data

To explore the importance of mosquito data (mosquito diversity and vector abundance) in explaining the spatial pattern of human infections, we re-ran the models including only the subset of the geographic locations for which we had mosquito data. Overall, the saturated models (which included all covariates listed in Table 3) for this subset of locations were able to explain more of the variation in human infections than the saturated models for the whole state, across all spatial scales. However, singly removing the mosquito variables from the saturated model only reduced the percent deviance explained by 0.1% for the two finer scale analyses and 0.3-0.4% for the coarsest scale. At all scales, the saturated model was stronger (lower AIC) without the mosquito variables (Table 3).

Discussion

Ross River virus is Australia's most epidemiologically important mosquito-borne pathogen, and its transmission is controlled by a variety of biotic and abiotic factors [8]. This study identified previously unrecognized determinants (i.e. urban score, density of caravan parks and reservoir host biodiversity) of RRV distribution in South Australia which has important implications for public health education and disease reduction strategies.

Traditionally, epidemiological studies analyze disease patterns with data aggregated at a single spatial scale. While such single-scale analyses are valuable, they paint an incomplete picture; important disease processes may be either overlooked or overstated in their importance. Our

finding that the spatial scale of an analysis influences the importance (and sometimes the direction of influence) of the potential drivers of disease, challenges the classic single spatial scale approach.

Other studies have found biotic variables to be more important at finer scale and abiotic (e.g. climate) variables to be more important at coarser scales [6,28–30]. Those of our variables that did show a pattern across scales followed this trend: biotic covariates relating to human exposure and vertebrate host abundance had greater effects at finer spatial scales, and abiotic factors of elevation and distance to the coast were more important at coarse spatial scales. This finding underlines the importance of multi-scale analyses generally, and the particular value of fine-scale analyses in identifying factors more easily modified by public health campaigns.

Biotic components

For the two finer scales, the expected number of infections (based on population size and age; see SM), was the most important determinant of the observed number of infections for a given area. At the coarsest scale of aggregation (region), the expected number of cases became negatively associated with the observed number of cases and urban score was the most important (negative) predictor. This counter-intuitive finding underlines the rural pattern of this infection; at finer scales, where some locations have few people to infect (e.g. fifteen suburbs have five or less people in them), and population variability is high, the abundance of people to infect drives the patterns of disease. But at the coarsest scale, where all areas have sufficient susceptible population and the relative population variability is lower, more cases occur in the rural areas despite their lower population.

This reversal displays that although RRV needs to infect a human to cause notifiable disease

(and therefore occurs in proportion to population), its ecology makes it, at coarser scales, a rural disease. The consistent negative association between urban score and RRV infections corresponds with the observation of higher RRV incidence in rural areas along the Murray River (Fig. 2) compared with the main metropolitan area of the state, Adelaide [20,31,32].

The number of caravan parks per person was found to be a positive predictor of disease at finer spatial scales, with no association found at a region spatial scale. Caravan parks are often located along water bodies that can breed mosquitoes and are associated with the type of outdoor activity that facilitates exposure to mosquitoes and their infections. Previous studies have implicated caravan parks as a risk factor for mosquito-borne diseases [20] but this hypothesis has not been tested until now.

Our inclusion of biodiversity variables relates to an ongoing debate whether host biodiversity reduces disease transmission (Dilution Effect Hypothesis; [26], increases transmission (Amplification effect) or if the effect of biodiversity depends on community composition, rather than diversity per se [33,34]. Bird species richness was negatively associated with RRV infections and was an important covariate among the finer spatial scales. Mammal richness had a weak but positive relationship with RRV disease incidence. Generally, mammals (especially marsupials) are more competent reservoirs for RRV than birds [8,10,35,36]. And while the main RRV vector species prefer to feed on mammals, their feeding preferences are flexible and they will feed on birds, especially if mammals are scarce [8,25,37]. Given current knowledge of reservoir capacity and vector feeding behavior, our findings support the hypothesis that birds may be acting as host “sinks” for RRV transmission. However, while these results are intriguing, in the absence of data on the abundance of animals in these taxa, it is impossible to determine

whether these associations are due to species richness per se, the community composition in biodiverse areas or the greater abundance which generally accompanies richness [38].

Socio-economic status (SES) was only found to be an important covariate at a fine spatial resolution, and had a negative association with the RRV notifications. Socio-economic status may be reflecting housing conditions, immunological conditions and/or human behaviors that influence human susceptibility to mosquitoes, infection or disease. Connections between SES and RRV infections have previously been made where education, awareness, and poor quality of housing have had links to increases in mosquito-borne disease risk [15,39]. Other studies, conducted using regional spatial scales, found no connection between SES and RRV infections [40]. Such discrepancies may result from including different covariates in the analyses or analyzing single-spatial scales.

Abiotic components

Of the abiotic environmental covariates tested, elevation was found to have a negative association with disease prevalence and had the strongest independent relationship with RRV infections across all spatial resolutions. Elevation has a well-known relationship with rainfall, temperature, and vector flight capabilities [41], though these correlations were not particularly pronounced among our data (Tables S4 and S5). We suggest that changes in elevation can potentially decrease vector mosquito breeding habitat, exposure to competent hosts, and reduced RRV exposure to human populations. The distance of an area from the Murray River was an important saturated model covariate at each spatial scale. The distance of an area from the Murray River had a relatively small effect on the number of RRV infections; the association was neutral at the finest spatial scale and negative at coarser scales. This result supports previous

investigations that showed elevated mosquito-borne disease infection rates in areas close to rivers [42,43]. The Murray River and its surrounds are often implicated as mosquito breeding habitat, habitat for RRV hosts, and as a recreational area for humans, all of which could contribute to the elevated RRV incidence there [20,24,31,32,44].

The distance between a site and the coast was also seen to be an important positive covariate at the suburb and regional spatial resolutions, with more disease further from the coast. This constrasts what is commonly seen elsewhere, where coastal areas are often associated with greater RRV infections. Coastal locations have previously been reported to have greater RRV infections than that of inland areas due primarily to increased tidal-driven mosquito breeding habitat [12,45,46]. Our analyses reveal that the distribution of RRV infections in South Australia follows the Murray River, rather than the coast.

The difference in biotic and abiotic drivers of RRV among various spatial scales demonstrates the complexity in transmission ecology and epidemiology, and the need for multi-spatial scale analyses when formulating mosquito intervention programs [4,6]. Furthermore, despite the novel inclusion of variables across all components of arboviral disease ecology, the spatial autocorrelation of the model residuals at the finer spatial scales (Table S6) indicates that some driver of the spatial pattern of RRV disease is still missing from this analysis. Intervention programs may not be fully effective until all drivers of infection are revealed.

Value of mosquito variables

Mosquito abundance and diversity did not improve the explanatory power of the models presented here, which contrasts previous epidemiological studies of RRV [18,45,47,48].

However, other studies often attempt to identify the conditions preceding an outbreak [45,47–49]. Such predictive models, while invaluable for public health, do not explain the consistent spatial patterns of the disease. Vector mosquito populations are clearly indicative of an outbreak but they appear to be redundant indicators of other biotic and abiotic variables when included in an ecological spatial analysis such as this one.

Conclusions

The analyses presented here not only advance our ability to explain patterns of RRV infection but elucidate the shifting importance of covariates across spatial resolutions. These models would be improved by including a temporal analysis to provide a complete multi-scale spatio-temporal advanced warning system for RRV. Such spatial and temporal eco-epidemiology studies are vital for disentangling the complex web of drivers of mosquito-borne diseases. By identifying specific factors associated with disease, these studies provide public health managers with information that can guide disease prevention programs. For example, from this analysis, we can recommend targeting public education materials to rural caravan parks, in particular suburbs, where people may engage in outdoor recreation behaviors that could expose them to mosquito-borne disease.

Unfortunately, epidemiological studies generally focus on a single spatial scale which can overlook important variables or over-emphasize the importance of variables, and lead to erroneous conclusions [4,6]. As a result, suboptimal outbreak responses may occur due to an incomplete understanding of disease ecology [50]. This study begins to address the multi-scale analysis void in the literature by evaluating the relative importance of major contributors of RRV disease ecology at distinct spatial scales. The conclusions presented here apply to the wider

vector-borne disease community and, if heeded, will fundamentally improve understanding and practice in the fields of public health, disease ecology and epidemiology.

Acknowledgements

The authors thank Dr. Andrew Flies for advice and edits on drafts and Dr, Barry Brook for statistical advice.

Works Cited

1. World Health Organization. A global brief on vector-borne diseases. Geneva, Switzerland; **2014**.

2. Reiner RC, Perkins TA, Barker CM, et al. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010. J R Soc Interface **2013**; 10(81):1–13.

3. Susser M, Susser E. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. Am J Public Health **1996**; 86(5):674–677.

4. Khormi HM, Kumar L. The importance of appropriate temporal and spatial scales for dengue fever control and management. Sci Total Environ **2012**; 430:144–149.

5. Tian H, Yan C, Xu L, Büntgen U, Stenseth NC, Zhang Z. Scale-dependent climatic drivers of human epidemics in ancient China. PNAS **2017**; doi:201706470.

6. Cohen JM, Civitello DJ, Brace AJ, et al. Spatial scale modulates the strength of ecological processes driving disease distributions. PNAS **2016**; 113(24):E3359-64.

7. Australian Government Department of Health. National Notifiable Diseases Surveillance System. Available from: http://www9.health.gov.au/cda/source/rpt_4_sel.cfm. Accessed 9 April, 2017.

8. Harley D, Sleigh A, Ritchie S. Ross River virus transmission, infection, and disease: a cross-disciplinary review. Clin Microbiol Rev **2001**; 14(4):909–932.

9. Claflin SB, Webb CE. Ross River virus : many vectors and unusual hosts make for an

unpredictable pathogen. PLOS Pathog **2015**; 11(9):1–5.

10. Russell RC. Ross River virus: ecology and distribution. Annu Rev Entomol **2002**; 47:1–31.

11. Kelly-hope LA, Purdie DM. Ross River virus disease in Australia , 1886 – 1998 , with analysis of risk factors associated with outbreaks. J Med Entomol **2004**; 41(2):133–150.

12. Gatton ML, Kelly-hope LA, Kay BH, Ryan PA. Spatial-temporal analyis of Ross River virus disease patterns in Queensland, Australia. Am J Trop Med Hyg **2004**; 71(5):629–635.

13. Gatton ML, Kay BH, Ryan PA. Environmental predictors of Ross River virus disease outbreaks in Queensland, Australia. Am J Trop Med Hyg **2005**; 72(6):792–799.

14. Vally H, Peel M, Hanigan I. Geographic Information Systems used to describe the link between the risk of Ross River virus infection and proximity to the Leschenault estuary, WA. Aust New Zeal J Public Heal **2012**; 36:229–235.

15. Hu W, Tong S, Mengersen K, Oldenburg B. Exploratory spatial analysis of social and environmental factors associated with the incidence of Ross River virus in Brisbane, Australia. Am J Trop Med Hyg **2007**; 76(5):814–819.

16. Ryan PA, Do K, Kay BH. Spatial and temporal analysis of Ross River virus disease patterns at Maroochy Shire, Australia : association between human morbidity and mosquito (Diptera : Culicidae) abundance. J Med Entomol **1999**; 36(4):515–521.

17. Ryan PA, Alsemgeest D, Gatton ML. Ross River virus disease clusters and spatial

relationship with mosquito biting exposure in Redland Shire, southern Queensland, Australia. J Med Entomol **2006**; 43(5):1042–1059.

18. Ng V, Dear K, Harley D, McMichael A. Analysis and prediction of Ross River virus transmission in New South Wales, Australia. Vector-Borne Zoonotic Dis **2014**; 14(6):422–438.

19. Australian Bureau of Statistics. Census of Population and Housing Table Builder. Available from: http://www.abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder. Accessed 21 Oct 2015

20. Flies EJ, Williams CR, Weinstein P, Anderson SJ. Improving public health intervention for mosquito-borne disease: the value of geovisualization using source of infection and LandScan data. Epidemiol Infect **2016**; doi: 10.1017/S09502688.

21. Harley D, Ritchie S, Bain C, Sleigh AC. Risks for Ross River virus disease in tropical Australia. Int J Epidemiol; **2005**; 34(3):548–555.

22. Pesaresi M, Huadong G, Blaes X, et al. A Global Human Settlement Layer From Optical HR / VHR RS Data : Concept and First Results. IEEE J Sel Top Appl Earth Obs Remote Sens **2013**; 6(5):2102–2131.

23. Williams CR, Kokkinn MJ, Snell AE, Fricker SR, Crossfield EL. Mosquitoes (Diptera: Culicidae) in metropolitan Adelaide, South Australia. Trans R Soc South Aust **2001**; 125:115–121.

24. Flies EJ, Toi C, Weinstein P, Doggett SL, Williams CR. Converting mosquito surveillance

to arbovirus surveillance with honey-baited nucleic acid preservation cards. Vector Borne Zoonotic Dis **2015**; 15(7):397–403.

25. Flies EJ, Flies AS, Fricker SR, Weinstein P, Williams CR. Regional comparison of mosquito bloodmeals in South Australia: implications for Ross River virus ecology. J Med Entomol **2016**; 53(4):902-910.

26. Ostfeld RS, Keesing F. The function of biodiversity in the ecology of vector-borne zoonotic diseases. Can J Zool **2000**; 78:2061–2078.

27. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available from: http://www.r-project.org/

28. Wiens JA. Spatial scaling in ecology. Funct Ecol **1989**; 3(4):385–397.

29. Levin SA. The problem of pattern and scale in ecology. Ecology **1992**; 73(6):1943–1967.

30. Chave J. The problem of pattern and scale in ecology : what have we learned in 20 years ? Ecol Lett. **2013**; Suppl 1:4-16.

31. Weinstein P, Worswick D, Macintyre A, Cameron S. Human sentinels for arbovirus surveillance and regional risk classification in South Australia. Med J Aust **1994**; 160(8):494–499.

32. Horwood CM, Bi P. The incidence of Ross River virus disease in South Australia, 1992 to 2003. Commun Dis Intell **2005**; 29(3):291–296.

33. LoGiudice K, Duerr STK, Newhouse MJ, Schmidt KA, Killilea ME, Ostfeld RS. Impact of host community composition on Lyme disease risk. Ecology **2008**; 89(10):2841–2849.

34. Salkeld DJ, Padgett KA, Jones JH. A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. Ecol Lett **2013**; 16(5):679-686.

35. Marshall ID, Miles JAR. Ross River virus and epidemic polyarthritis. Curr Top Vector Res **1984**; 2:31–56.

36. Koolhof IS, Carver S. Epidemic host community contribution to mosquito-borne disease transmission: Ross River virus. Epidemiol Infect **2017**; 145(4):656–666.

37. Kay BH, Boreham PFL, Fanning ID. Host-feeding patterns of Culex annulirostris and other mosquitoes (Diptera: Culicidae) at Charleville, southwestern Queensland, Australia. J Med Entomol **1985**; 22(5):529–535.

38. LoGiudice K, Ostfeld RS, Schmidt KA, Keesing F. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. PNAS **2003**; 100(2):567–571.

39. Weinstein P, Cameron S. A houseboat outbreak of epidemic polyarthritis. Med J Aust **1991**; 155:721–722.

40. Hu W, Clements A, Williams G, Tong S, Mengersen K. Bayesian spatiotemporal analysis of socio-ecologic drivers of Ross River virus transmission in Queensland, Australia. Am J Trop Med Hyg **2010**; 83(3):722–728.

41. Verdonschot PFM, Besse-Lototskaya AA. Flight distance of mosquitoes (Culicidae): A metadata analysis to support the management of barrier zones around rewetted and newly

constructed wetlands. Limnol - Ecol Manag Inl Waters **2014**; 45:69–79.

42. Bi P, Hiller JE, Cameron AS, Zhang Y, Givney R. Climate variability and Ross River virus infections in Riverland, South Australia, 1992–2004. Epidemiol Infect **2009**; 137(10):1486–1493.

43. Jardine A, Neville PJ, Lindsay MDA. Proximity to mosquito breeding habitat and Ross River virus risk in the Peel Region of Western Australia. Vector-Borne Zoonotic Dis **2015**; 15(2):141–146.

44. Mudge PR, Lim RSH, Moore B, Radford AJ. Epidemic polyarthritis in South Australia 1979-1980. Med J Aust **1980**; 2(11):626–627.

45. Jacups SP, Whelan PI, Markey PG, Cleland SJ, Williamson GJ, Currie BJ. Predictive indicators for Ross River virus infection in the Darwin area of tropical northern Australia, using long-term mosquito trapping data. Trop Med Int Heal **2008**; 13(7):943–952.

46. Tong S, Bi P, Donald K, McMichael AJ. Climate variability and Ross River virus transmission. J Epidemiol Community Heal **2002**; 56:617–621.

47. Hu W, Tong S, Mengersen K, Oldenburg B. Rainfall , mosquito density and the transmission of Ross River virus : A time-series forecasting model. Ecol Modell **2006**; 196:505–514.

48. Woodruff RE, Guest CS, Garner MG, Becker N, Lindsay M. Early warning of Ross River virus epidemics combining surveillance data on climate and mosquitoes. Epidemiology **2006**; 17(5):569–575.

49. Woodruff RE, Guest CS, Garner MG, et al. Predicting Ross River virus epidemics from regional weather data. Epidemiology **2002**; 13(4):384–393.

50. Brady OJ, Smith DL, Scott TW, Hay SI. Dengue disease outbreak definitions are implicitly variable. Epidemics **2015;** 11:92–102.

Table 1. State geographical divisions, sample size and area

Code	Name	Total N	N for mosquito data subset	Median area (km^2)	Standard dev. (km^2)
SSC	Suburb	858	219	6.9	12,815.9
POA	Post code	319	100	39.4	21,961.2
SLA	Region	127	51	193.8	43,693.6

Table 2. All covariates a priori believed to be important in determining the distribution of cases of Ross River virus infection, with a description of the variable, and the predicted relationship between covariate and Ross River virus infection counts. Citations to support the a priori hypothesis and source of data are available in Table S2. Neg = negative, Pos = positive.

Category	Covariate abbreviation	Description	A priori hypothesized relationship
Humans	SES	Average index of relative socioeconomic disadvantage score for area (SEIFA)	Neg
	Caravans	Number of caravan parks per person for the area	Pos
	UrbanScore	Average global human settlement "urbanness" score	Neg
Habitat	DistToCoast	Distance from centroid to coast	Neg
	DistToMurray	Distance from centroid to Murray River	Neg
	NDVI	Average 16-day average normalized difference in vegetation index (Jan 1	Pos

		2011)	
	Elevation	Average elevation (from SRTM-derived 1 second digital elevation model)	NA
Reservoirs	Bird	Bird species richness	Neg
	Mamm	Mammal species richness	Pos
	Marsup	Marsupial species richness	Pos
Climate	Rain	Average 30 year 6 month (Oct-April) average rainfall	Pos
	MinT	Average 30 year 6 month (Oct-April) average minimum temperature	Pos
	RelHum	Average 30 year 6 month (Oct-April) average relative humidity	Pos
	MaxT	Average 30 year 6 month (Oct-April) average maximum temperature	Pos
Vectors	Vectors	Maximum number of vector mosquitoes	Pos
	NonVect	Maximum number of non-vector	Neg

		mosquitoes	
	Shannon	Mosquito Shannon diversity	Pos
Expected	ExpectedCases	Expected number of cases	Pos

Table 3. Relationships and importance of individual covariates in explaining the Ross river virus infection count per area across three spatial scales. Vector abundance and (Shannon) diversity of mosquitoes were only available for the subset of locations for which mosquito data was available, thus AIC and % deviance explained can only be compared within that analysis, not between that analysis and the state-wide analysis.

	Suburbs (n = 858, area = 6.9 km^2)				Post Codes (n = 319, area = 39.4 km2)				Regions (n = 127, area = 193.8 km2)			
Covariate	Single model coeff	Single model SE	Δ AICc when removed	Δ %Dev when removed	Single model coeff	Single model SE	Δ AICc when removed	Δ %Dev when removed	Single model coeff	Single model SE	Δ AICc when removed	Δ %Dev when removed
Saturated model % deviance explained				40.2				53.0				45.3
ExpectedCases	0.57	0.11	325.63	20.5	0.30	0.13	132.61	19.8	-0.33	0.14	14.62	8.5
UrbanScore	-0.55	0.12	144.56	6.8	-0.52	0.12	59.64	8.8	-0.64	0.13	42.37	19.2
Caravans	0.52	0.12	41.92	3.1	0.19	0.13	5.08	0.9	0.21	0.15	-0.27	1.3
DistToMurray	0.00	0.12	30.8	2.3	-0.03	0.13	4.43	1.0	-0.18	0.14	3.19	2.6
Bird	-0.63	0.11	29.13	2.2	-0.55	0.11	14.50	2.5	-0.36	0.15	-2.04	0.4

Elevation	-0.35	0.11	21.78	1.5	-0.51	0.11	52.01	6.6	-0.26	0.14	22.51	7.7
SES	-0.85	0.09	4.81	0.5	-0.64	0.10	-2.11	0.0	-0.52	0.12	-1.74	0.3
Mamm	0.36	0.10	6	0.3	0.42	0.12	-1.19	0.2	0.33	0.14	1.34	1.2
MaxT	0.45	0.11	-1.41	0.2	0.38	0.11	0.19	0.3	0.30	0.14	0.90	1.1
DistToCoast	0.70	0.10	-3.51	0	0.51	0.11	0.93	0.5	0.44	0.13	5.18	3.4
NDVI	-0.33	0.13	-1.91	0	-0.29	0.13	-1.52	0.1	-0.16	0.15	-0.14	1.1
MinT	0.06	0.12	-2.52	0	0.09	0.13	0.31	0.3	-0.08	0.15	-1.72	0.3
Mosquito data subset (n = 219, 100, 51, respectively)												
Saturated model % deviance explained				68.2				79.8				66.1
Vector*	1.29	0.18	-1.009	0.1	1.12	0.16	-2.5249	0.1	1.05	0.22	-4.84	0.3
Shannon*	0.61	0.27	-1.3593	0.1	0.68	0.19	-2.4727	0.1	0.80	0.20	-3.02	0.4

Figure 1**.** Components to disease ecology and how covariates fit within those components. Covariates with dashed boxes may represent aspects of multiple disease ecology components. Details of these covariates are located in Table S2.

Figure 1. State-wide Ross River virus infection incidence according to patient-reported source of infection (n = 858) and the abundance of vector mosquitoes (points on inset map), where such data were available.

Figure 3. Changing importance of covariates across scales, based on the relative rank of the change in percent deviance explained (Δ %Dev Exp) when each variable is singly removed from the saturated model. Scale becomes more coarse from left to right, 10 = most important, 1 = least important.

Figure 1.

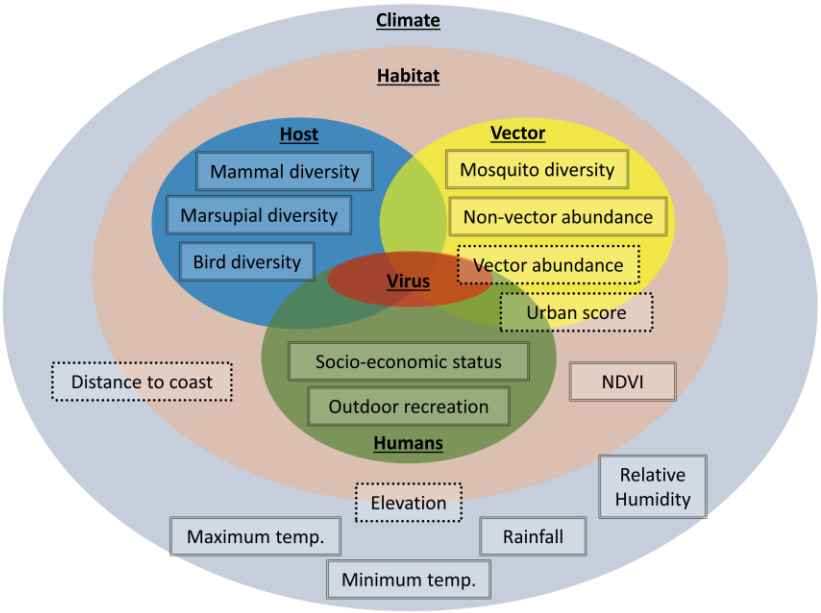


Figure 2.

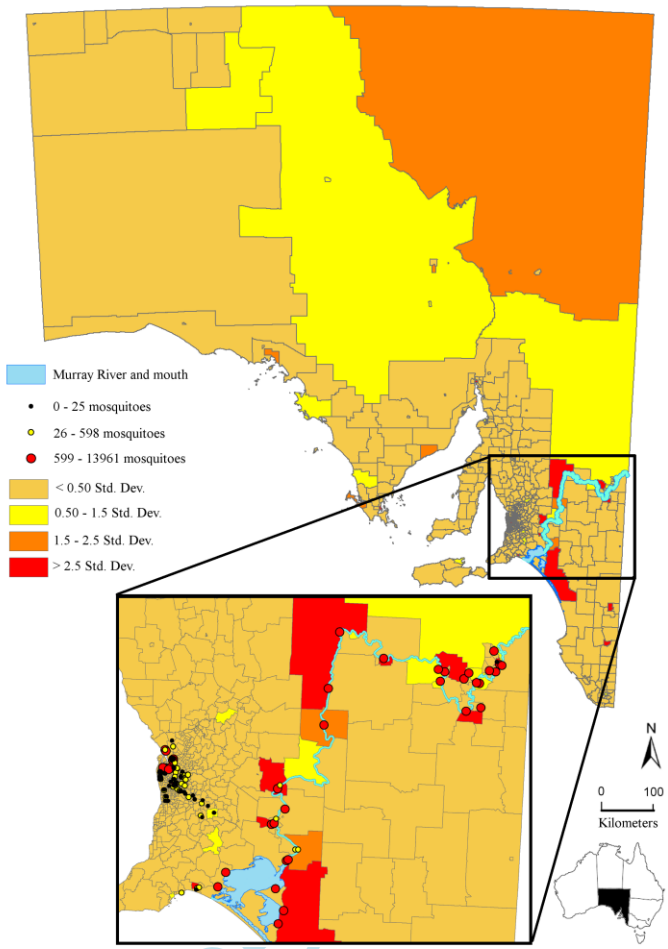
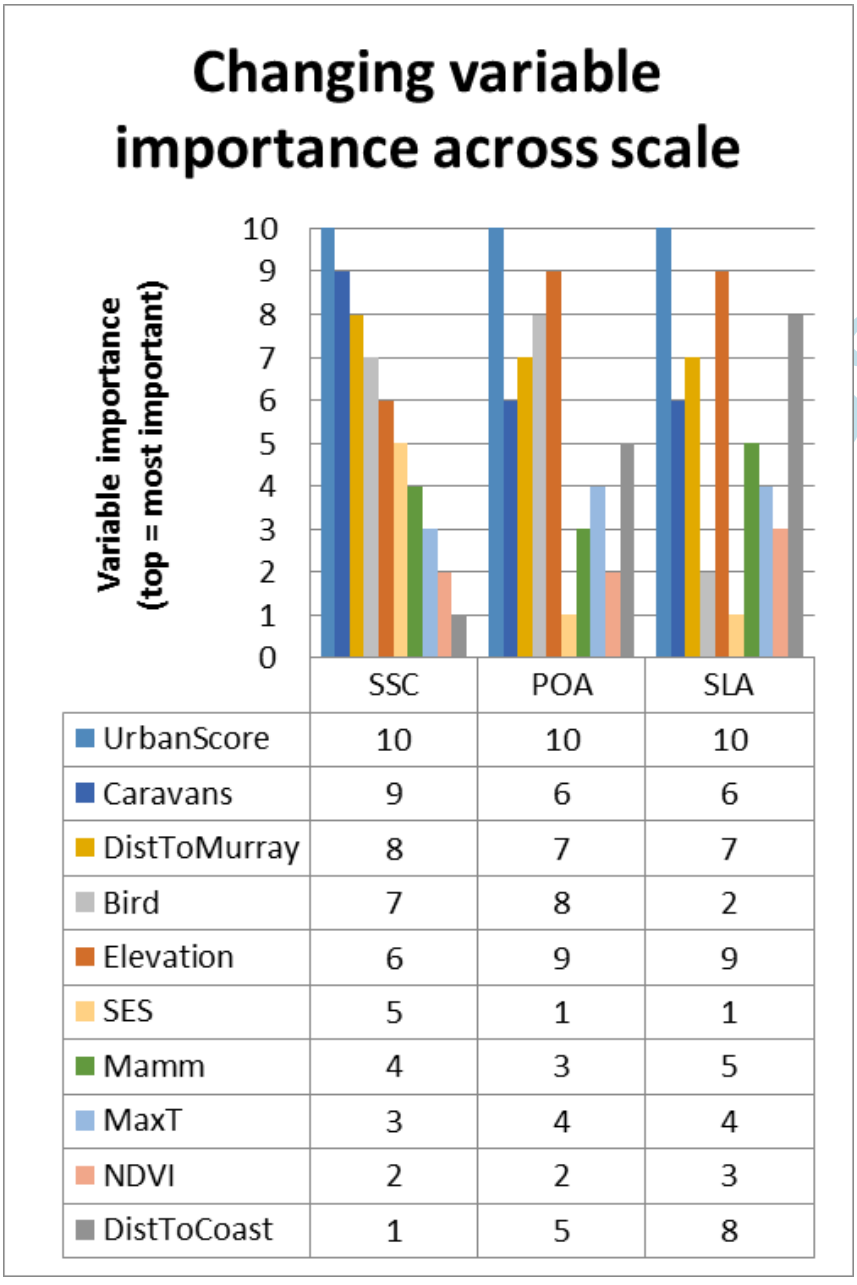


Figure 3.



	SSC	POA	SLA
UrbanScore	10	10	10
Caravans	9	6	6
DistToMurray	8	7	7
Bird	7	8	2
Elevation	6	9	9
SES	5	1	1
Mamm	4	3	5
MaxT	3	4	4
NDVI	2	2	3
DistToCoast	1	5	8